Proc. Natl. Acad. Sci. USA
Vol. 90, pp. 11608–11612, December 1993
Evolution

Evolution of nuclear ribosomal RNAs in kinetoplastid protozoa: Perspectives on the age and origins of parasitism

(RNA editing/mitochondrial rRNA/*Bodo caudatus*/Trypanosomatidae)

ANA PAULA FERNANDES, KIMBERLYN NELSON*, AND STEPHEN M. BEVERLEY†

Department of Biological Chemistry and Molecular Pharmacology, Harvard Medical School, Boston, MA 02115

Communicated by Walter Gilbert, September 7, 1993 (received for review July 16, 1993)

ABSTRACT **Molecular evolutionary relationships within the protozoan order Kinetoplastida were deduced from comparisons of the nuclear small and large subunit ribosomal RNA (rRNA) gene sequences. These studies show that relationships among the trypanosomatid protozoans differ from those previously proposed from studies of organismal characteristics or mitochondrial rRNAs. The genera *Leishmania, Endotrypanum, Leptomonas*, and *Crithidia* form a closely related group, which shows progressively more distant relationships to *Phytomonas* and *Blastocrithidia, Trypanosoma cruzi*, and lastly *Trypanosoma brucei*. The rooting of the trypanosomatid tree was accomplished by using *Bodo caudatus* (family Bodonidae) as an outgroup, a status confirmed by molecular comparisons with other eukaryotes. The nuclear rRNA tree agrees well with data obtained from comparisons of other nuclear genes. Differences with the proposed mitochondrial rRNA tree probably reflect the lack of a suitable outgroup for this tree, as the topologies are otherwise similar. Small subunit rRNA divergences within the trypanosomatids are large, approaching those among plants and animals, which underscores the evolutionary antiquity of the group. Analysis of the distribution of different parasitic life-styles of these species in conjunction with a probable timing of evolutionary divergences suggests that vertebrate parasitism arose multiple times in the trypanosomatids.**

The protist order Kinetoplastida is a cosmopolitan group of flagellates containing prominent groups that parasitize virtually every major group of eukaryotic organisms, including other protozoans (1, 2). Kinetoplastids are distinguishable from other protozoa by the presence of kinetoplast DNA, a different type of mitochondrial DNA consisting of maxi- and mini-circle DNAs located in the single mitochondrion near the basal body of the flagellum (1, 3). In addition, these organisms exhibit unusual molecular phenomena such as antigenic variation, trans-splicing, and RNA editing (3–5). Kinetoplastids exhibit three different life-styles: free-living, monogenetic parasitism involving a single host (usually an invertebrate), and digenetic parasitism alternating between invertebrate and vertebrate or plant hosts. The family Bodonidae contains both free-living and monogenetic taxa, whereas the family Trypanosomatidae is a diverse group of monogenetic and digenetic taxa (1, 2). In humans and animals *Trypanosoma* and *Leishmania* cause severe diseases, including sleeping sickness, Chagas disease, and kala-azar, while *Phytomonas* are pathogens of plants.

Previous studies of evolutionary relationships within the Trypanosomatidae emphasized traits derived from morphology or consideration of the parasitic life-style. Many workers favored a model where the digenetic genera such as *Endotrypanum, Leishmania*, and *Trypanosoma* originated from monogenetic, *Leptomonas*-like ancestors (6, 7). However, the number of independent origins of digenetic parasitism has been estimated variously as from 1 to 5 (6–9). An alternative model suggests that digenetic parasites originated from free-living invaders and secondarily gave rise to monogenetic parasites (10). Thus, an unambiguous view of both parasite relationships and the origins of vertebrate parasitism has not emerged.

Recently molecular evolutionary methods have been applied to the study of trypanosomatid relationships. Comparisons of mitochondrial (kinetoplast) rRNAs suggested that the genera *Crithidia, Leptomonas, Leishmania*, and *Trypanosoma* diverged successively, with the digenetic life-style originating once before the separation of *Leishmania* (11). However, analysis of nuclear small subunit ribosomal RNAs (SSU rRNAs) suggested a different evolutionary tree, in which *Trypanosoma brucei* and *Trypanosoma cruzi* diverge before the lineage leading to *Crithidia* and *Leishmania*, which are closely related (Figs. 1–4; refs. 12–14). The implications of these various topologies to models concerning the age and origins of parasitism are profound. Significantly, the mitochondrial tree lacked an independent outgroup required to root the tree, which cast doubts about the direction of evolution. Were the root placed on the *T. brucei* lineage, the topologies of the mitochondrial and nuclear trees would be quite similar. In this study we report on the sequence of the SSU and large subunit (LSU) rRNAs from a close outgroup relative of trypanosomatids, *Bodo caudatus* of the Kinetoplastid family Bodonidae (15), as well as additional sequences for monogenetic (*Leptomonas, Blastocrithidia*) and digenetic (*Endotrypanum, Leishmania, Phytomonas*) trypanosomatids. These data permit an unambiguous rooting of the nuclear evolutionary tree, which confirms the view that *T. brucei* represents the most ancient divergence within the trypanosomatids. We consider this tree to accurately depict phylogenetic relationships within the Trypanosomatidae and use this perspective to view the age and origins of vertebrate parasitism in this important group.

MATERIALS AND METHODS

Parasite Strains and DNAs. The strains were as follows: *B. caudatus* (a clonal derivative obtained from S. Hajduk, University of Alabama, Birmingham; ref. 16), *T. cruzi* (Peru strain; from D. McMahon-Pratt, Yale University Medical School), *Endotrypanum monterogei* (strain LV88; from M. Chance, Liverpool School of Tropical Medicine), *Leptomonas* sp. (from L. Simpson, University of California, Los Angeles; ref. 11), *Leishmania major* and *L. donovani* (strains LT252 and HU-3, this laboratory), *Phytomonas* sp. (isolated from *Euphorbia hyssopifolia*; from I. Roitman, Universidade de Brasilia, ref. 17) and *Blastocrithidia culicis* (ATCC 30268,

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Abbreviations: SSU rRNA, small subunit ribosomal RNA; LSU rRNA, large subunit ribosomal RNA.

*Present address: Institute of Molecular Evolutionary Genetics, Department of Biology, Pennsylvania State University, University Park, PA 16802.

†To whom reprint requests should be addressed.

from L. Landweber, Harvard University). *Bodo* cells were separated from the feeder bacteria by velocity and density-gradient centrifugation (16). DNAs were isolated by preparation of chromosomes in agarose plugs (*Blastocrithidia, Endotrypanum*, and *Leishmania*; ref. 18), banding in CsCl gradients (*Bodo, Leptomonas*; ref. 19), or phenol-chloroform extraction (*T. cruzi, Phytomonas*; ref. 20).

PCR Amplification and DNA Sequencing. Nuclear rRNA genes were amplified with oligonucleotide primers chosen from conserved sequences evident in published sequences of *T. brucei, Leishmania donovani*, and *Crithidia fasciculata* (21–25). The entire SSU rRNA gene was amplified by using the primers situated at the extreme 5′ and 3′ termini of the rRNA coding region. Primers for a 1-kb segment of the LSU rRNA were xba-CLS378 (5′-GGGTCTAGAGTAGGAAGAC-CGATAGC) and kpn-CLS1373 (5′-GTGGTACCGGTGGAT-TCGGTTGGTGAG). PCR amplification (26) was performed by 30 cycles, consisting of 1 min at 94°C, 2 min at 55°C, and 3 min at 72°C. Amplified DNAs were digested with appropriate enzymes, purified on low-melting-temperature agarose gels, and cloned into appropriately digested M13mp18 and M13mp19 vectors (27). Single-stranded DNAs were prepared for several clones from each orientation and sequenced individually or after pooling. Both strands were sequenced by the dideoxynucleotide chain-termination method by using either the M13 sequencing primers or primers complementary to conserved regions of the gene. Accession numbers for SSU sequences are as follows: *B. caudatus*, X53910; *T. cruzi* Peru, X53917; *E. monterogei*, X53911; and *Leptomonas* sp., X53914. Accession numbers for LSU sequences are as follows: *B. caudatus*, L19405 and L19406; *T. cruzi* Peru, L19411; *B. culicis*, L19404; *Phytomonas* sp., L19410; *L. donovani*, L19408; *E. monterogei*, L19407; *Leptomonas* sp., L19409. Aligned sequences and filters in computer-readable format are available from the authors.

Sequence Alignment and Filters. Two different SSU alignments were examined. Alignment 1 was provided by M. Sogin, Marine Biological Laboratories, Woods Hole, MA and is based upon previous alignments used by this group. Alignment 2 was derived from published alignments (28) that included *T. brucei, Crithidia*, and *L. donovani* sequences. Sequences from the remaining taxa obtained in this work were added initially by inspection and refined by using a variety of alignment algorithms, as well as consideration of the secondary structure where possible. LSU alignments were prepared similarly. Because meaningful alignments for all regions of the LSU and SSU sequences could not be obtained among all taxa, the data sets required filtering to retain only positions that were unambiguously homologous. For alignment 1, we used filters provided by Sogin; from a total of 2540 total positions, 1609 and 2199 positions were retained with or without *Euglena* as an outgroup, respectively. A computer-based filter described below was also used to select objectively alignable regions. For SSU rRNA alignment 1, this retained 1329 positions with and 2000 positions without *Euglena* as an outgroup. For SSU rRNA alignment 2, from a total of 2664 positions 1264 and 2074 positions were retained with and without *Euglena* as an outgroup. For the LSU rRNA alignment (containing only kinetoplastid species), from a total of 1159 positions 812 were retained. Distance comparisons among sequences were calculated by using Sogin's structural similarity measure (29). Evolutionary trees were calculated and evaluated by parsimony methods with the PAUP 3.1.1 program (30). Bootstrap analysis was performed by using 500 replicas.

A computer-based filter for selecting "alignable" positions was developed, starting with an aligned matrix of X sequences containing Y positions. Beginning at position 1, a window of the first N nucleotides was selected, and all pairwise comparisons were calculated among the X sequences. If the minimum match obtained in these pairwise comparisons were at least A, all N positions were considered to be sufficiently related for retention. Then, the pairwise comparisons were repeated successively with windows beginning at aligned position 2, 3, . . . $(Y - N + 1)$ to give the final filtered data set. Empirical evaluation suggested that a filter (A/N) of 4/7 yielded satisfactory data sets (as shown by the results obtained above with alignment 1).

RESULTS AND DISCUSSION

Sequences and Methodology. We used PCR amplification followed by molecular cloning to sequence the entire 2.2-kb SSU rRNA gene and a 1-kb segment of the LSU rRNA gene from *B. caudatus, T. cruzi* (Peru strain), *E. monterogei, Leptomonas* sp. (11), and *L. donovani* (LSU rRNA only); the LSU and SSU rRNA sequences of *C. fasciculata* and *T. brucei* have been reported, as have the SSU rRNA sequences for five species of *Leishmania* (14, 21–25, 31). LSU sequences were additionally obtained for isolates of *Phytomonas* and *B. culicus*. Multiple independent PCR clones were sequenced from each species, and some variation was evident. For example, within a species variation of 0.02% was seen among different PCR-generated recombinants of the SSU rRNA gene. This heterogeneity arose from both nucleotide substitutions and single-base insertions and deletions and, presumably, includes variation among individual ribosomal genes as well as errors due to PCR amplification. Variation among individual ribosomal genes is known to occur within several trypanosomatid species (32). Regardless of its source, the experimental intraspecific variability is much lower than that among the species discussed in this work and does not significantly affect the conclusions drawn herein.

For *B. caudatus*, two different PCR products comprising LSU rRNA sequences were obtained, differing 1.5% in nucleotide sequence and in length by 25 nt arising from 18 small insertions. The LSU and SSU sequences from the Peru strain of *T. cruzi* differed from those reported (12, 32) for other strains. The SSU rRNA sequence of a Mexican strain showed a difference of 0.3%, whereas the LSU sequence of the Y strain showed a difference of 1%. This high degree of intraspecific divergence within *T. cruzi* is consistent with the extremely high divergence noted in isoenzyme loci (33). Previous work (14, 31, 34) has shown that the SSU rRNA genes of *Leishmania* species are closely related, and they will not be considered separately here.

Evolutionary Trees. Comparisons of the genetic distances among SSU rRNA sequences showed that *Leishmania, Endotrypanum, Crithidia*, and *Leptomonas* were most similar and formed a group that was progressively less similar to *T. cruzi, T. brucei, Bodo*, and *Euglena* (Table 1). The same pattern was evident in comparisons of LSU rRNA sequences from these species (Table 1).

Molecular trees were constructed by using maximum-parsimony methods (Figs. 1–3). We first asked whether *Bodo* was placed outside the Trypanosomatidae as expected (15), using *Euglena* as the outgroup. Maximum-parsimony trees with both SSU rRNA alignments 1 or 2 placed *Bodo* as an outgroup (as did the 15 next best trees; results with alignment 2 are shown in Fig. 1). Bootstrap replica analysis supported this placement in 99% and 81% of the replicas for alignments 1 and 2, respectively (Fig. 1). Evolutionary parsimony analysis of either alignment 1 or 2 did not support the placement of *Bodo* as an outgroup. This result may be some peculiarity arising from *Euglena*, as the use of other taxa as outgroups (such as *Dictyostelium, Zea mays*, human, *Chlamydomonas*) placed *Bodo* outside the Trypanosomatidae (data not shown). This result confirms the outgroup status of *Bodo* by both organismal and molecular criteria.

To view relationships within the Trypanosomatidae, trees were constructed by using *Bodo* as the outgroup, as this

Table 1. rRNA comparisons among Kinetoplastida

	Leish.	*Endo.*	*Lept.*	*Crith.*	*Phyto.*	*Blasto.*	*T. cruzi*	*T. brucei*	*Bodo*	*Euglena*
Leishmania	0.997*	0.989	0.983	0.987	—	—	0.933	0.900	0.868	0.585
Endotrypanum	0.999	0	0.983	0.983	—	—	0.934	0.901	0.867	0.584
Leptomonas sp.	0.992	0.994	0	0.985	—	—	0.936	0.898	0.871	0.584
Crithidia	0.986	0.987	0.988	0	—	—	0.936	0.906	0.867	0.587
Phytomonas	0.978	0.979	0.978	0.967	0	—	—	—	—	—
Blastocrithidia	0.957	0.959	0.955	0.948	0.956	0	—	—	—	—
T. cruzi	0.975	0.975	0.971	0.967	0.976	0.950	0	0.908	0.872	0.585
T. brucei	0.949	0.950	0.947	0.943	0.95	0.934	0.945	0	0.868	0.592
Bodo†	0.937	0.937	0.937	0.936	0.939	0.918	0.947	0.920	0	0.577

Structural-similarity values were calculated as described (29) from the LSU or SSU (alignment 2) rRNA data sets by using a 4/7 filter criterion developed from kinetoplastid species only. SSU rRNA values are above the diagonal; LSU rRNA values are below the diagonal. A dash signifies that the comparison was not possible. Column-head abbreviations indicate the same genera that indicate rows.

*This value is the average SSU rRNA distance among five species of *Leishmania* (*L. major, L. donovani, L. panamaensis, L. amazonensis, L. tarentolae*).

†These values represent the average of values obtained with the two different *Bodo* LSU rRNA sequences (the separate values differ by <0.005).

permitted retention of more informative sites. The relationships obtained with the combined LSU + SSU (alignment 2) rRNA data set are shown in Fig. 2. The maximum-parsimony tree showed the trypanosomes as polyphyletic, with *T. brucei* diverging before *T. cruzi* (Fig. 2). The polyphyletic origin of trypanosomes was favored in the combined LSU and SSU rRNA alignment 2 data sets (77, 52, and 81% of bootstrap replicas; Figs. 1–3), and evolutionary parsimony analysis of the LSU alignment favored the polyphyletic origin. Hernandez and coworkers (12, 13) have also suggested a polyphyletic origin of trypanosomes, on the basis of comparisons of a limited number of SSU and smaller rRNAs. However, the SSU rRNA alignment 1 favored a monophyletic clustering of the trypanosomes (75% of bootstrap replicas), whereas evolutionary parsimony analysis of the two SSU alignments did not strongly favor either a mono- or polyphyletic origin. Differences arising from different molecules and sequence alignments are often taken to indicate uncertainty in the branching order. Although at this time the data favor a polyphyletic origin of trypanosomes, this conclusion should be considered tentative.

The molecular trees clustered the sand fly-borne digenetic parasites *Leishmania* and *Endotrypanum* together. This clustering was supported by bootstrap analysis of the combined data set and SSU rRNA alignments 1 and 2 (67, 82, and 91% respectively) and by evolutionary parsimony analysis. For the LSU rRNA we additionally sequenced *Phytomonas* (digenetic plant parasite) and *Blastocrithidia* (monogenetic); the molecular tree for these are shown in Fig. 3. Both of these taxa diverge from the *T. cruzi–Leishmania* lineage before the separation of *Crithidia, Leptomonas*, and *Endotrypanum*. A long branch length is inferred for the *Blastocrithidia* lineage, accounting for its larger divergence in distance comparisons (Table 1).

Nuclear vs. Mitochondrial (kinetoplast) rRNA Trees. The published mitochondrial rRNA tree (11) differs substantially from the nuclear rRNA tree presented here. However, the topology of the mitochondrial rRNA tree (obtained by evolutionary or standard parsimony methods; data not shown) is similar to that of the nuclear rRNA tree, differing primarily in locating the root either on the *Crithidia* lineage (mitochondrial tree) or *T. brucei* lineage (nuclear tree; refs. 13 and 14). The first rooting leads to trees for which branch lengths vary greatly among lineages, whereas the nuclear rooting shows more even rates of molecular evolution (data not shown). Because the mitochondrial data set lacked an outgroup, we attempted to obtain the *Bodo* mitochondrial rRNA sequence. We were unable to isolate this gene by PCR amplification, despite the use of a variety of template preparations (nuclear, extrachromosomal DNA; total RNA) and a wide selection of primers to regions conserved throughout the Trypanosomatidae; in all cases successful amplification was obtained with trypanosomatid templates and with all *Bodo* templates using nuclear rRNA primers. Similarly, Hadjuk *et al.* (16) were unable to identify mitochondrial rRNA genes in Southern blot analyses of *Bodo* DNA with *T. brucei* probes. In the absence of a proper outgroup for the mitochondrial data set, we

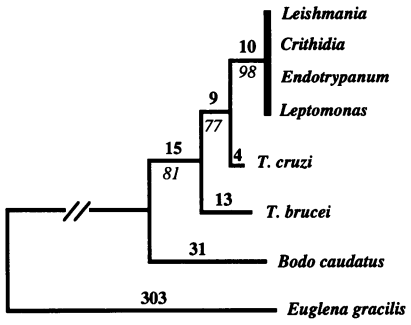

FIG. 1. Evolutionary relationships among kinetoplastids on the basis of SSU rRNA comparisons. SSU rRNA alignments (alignment 2) were filtered by using the 4/7 criterion described in text. All possible trees were examined by maximum parsimony. This tree is the most parsimonious and has a length of 393 with 31 phylogenetically informative sites; all of the 16 next best trees placed *Bodo* as the outgroup. Similar results were obtained with alignment 1. Numbers above each branch represent the calculated length, and numbers in italics below each branch represent the percentage of bootstrap replicas yielding this tree.

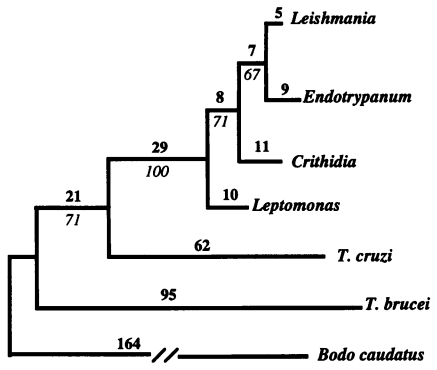

FIG. 2. Evolutionary relationships among trypanosomatids on the basis of SSU and LSU rRNA comparisons. LSU and SSU rRNA alignments (alignment 2) were filtered by using the 4/7 criterion as described in text; 2628 positions were retained. All possible trees were examined by maximum parsimony. The shortest tree is shown, with a length of 421 on the basis of 103 informative sites. The three next best trees similarly place *T. brucei* outside of *T. cruzi*. Numbers above each branch represent the calculated length, and numbers in italics below each branch represent the percentage of bootstrap replicas yielding this tree.

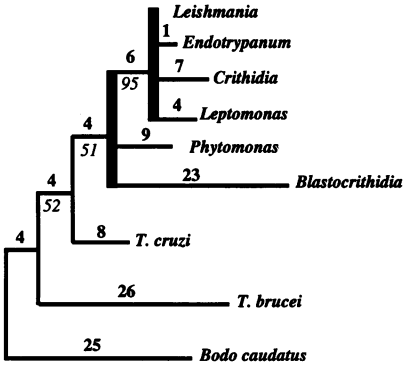

FIG. 3. Evolutionary relationships among trypanosomatids on the basis of LSU rRNA comparisons. The tree was calculated from the aligned LSU rRNA sequences, filtered by using the 4/7 criterion described in text. The shortest tree had a length of 119 on the basis of 21 informative sites; the tree shown has this length and collapses the *Blastocrithdia–Phytomonas* and *Leishmania–Endotrypanum–Crithidia–Leptomonas* lineages. Numbers above each branch represent the calculated length, and numbers in italics below each branch represent the percentage of bootstrap replicas yielding this tree.

presume that there is, in fact, no conflict between the mitochondrial and nuclear trees.

Other Molecular Comparisons Support the Nuclear rRNA Tree. The relationships shown in Figs. 1–3 are supported by many other molecular comparisons. For example, Southern blot studies with three different nuclear probes from *L. major* showed intense hybridization to *Crithidia, Endotrypanum*, and *Leptomonas* DNAs, weak hybridization to *T. cruzi* DNA, and faint hybridization or none at all to *T. brucei* or *Bodo* DNA (ref. 35; unpublished work). Association of *Crithidia* and *Leishmania* relative to the trypanosomes has been observed for numerous proteins, including dihydrofolate reductase, thymidylate synthase, trypanothione reductase, and mitochondrially encoded proteins (refs. 36–38; A. Fairlamb, personal communication; unpublished data). *Leishmania* and *Crithidia* both exhibit a very G + C-rich bias in the third-codon position, whereas *T. brucei* is not highly biased, and *T. cruzi* is intermediate (39–41). Chromosomal patterns obtained after electrophoretic separations by pulsed-field electrophoresis also show relationships consistent with these trees (unpublished data). Together, these findings confirm the close relationship of *Crithidia* and *Leishmania* and suggest that the rRNA relationships inferred in this work accurately reflect the relationships among trypanosomatid genes generally.

Origins of Parasitism. To view the evolutionary trends within the Trypanosomatidae, a tree summarizing the relationships deduced above (Fig. 4) is shown with each taxon labeled as free-living (F), monogenetic parasites of insects (M), or digenetic parasites of insects and vertebrates (D) or plants (P). This tree suggests that the origins of different parasitic life-styles are more complex than previously supposed and require multiple events. At least two different scenarios can be deduced from these data: (*i*) Parasitism of both insects and vertebrates is proposed to originate after the divergence of *Bodo* from the trypanosomatids. Then, vertebrate parasitism was lost after the divergence of the *Leishmania* and *T. cruzi* lineages and re-acquired in the common ancestor of *Leishmania* and *Endotrypanum*. Although loss of digenetic parasitism events are not often considered, the ease with which trypanosomatids can lose infectivity to vertebrates during routine laboratory culture suggests its potential occurrence in nature (10). The origins of plant parasitism in *Phytomonas* could have occurred either by a lateral shift from digenetic insect/vertebrate to insect/plant parasitism or by *de novo* acquisition of digenetic plant parasitism from the proposed monogenetic ancestor.

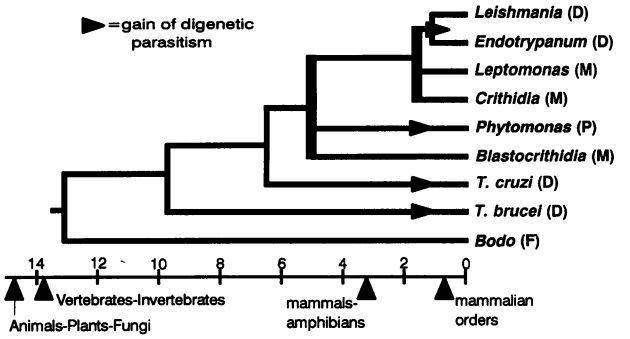

FIG. 4. Relationships and origins of parasitism within the trypanosomatidae. This tree represents a synthesis of results obtained for all taxa and sequences discussed in this work; vertical distances were for clarity only. The horizontal scale corresponds to the average percentage of SSU rRNA divergence among lineages, calculated from the structural similarity values in Table 1. Positions of *Blastocrithidia* and *Phytomonas* were estimated by comparing the LSU and SSU divergences of the other species and using this to estimate the expected SSU divergence. Horizontal arrowheads represent origins of digenetic parasitism, as discussed in text. D, digenetic vertebrate parasites; P, digenetic plant parasites: M, monogenetic parasites; F, free-living organisms. For comparisons among metazoans the following taxa were used: *Neurospora crassa, Saccharomyces cerevisae, Zamia pumila, Arabidopsis thaliana, Oryza sativa, Zea mays, Glycine max, Lycopersicon esculentum, Caenorhabditis elegans, Drosophila melanogaster, Tenebrio molitor, Eurypelma californica, Artemia salina, Xenopus laevis, Xenopus borealis*, rat, mouse, and human. Distances were estimated from structural-similarity comparisons (29) of the SSU rRNA sequence alignment (28) after using a filter criterion of 4/7. Vertical arrowheads represent the average value for the indicated comparisons.

(*ii*) Digenetic parasitism originated at least four different times from a monogenetic or even free-living ancestor: once each in the *T. brucei, T. cruzi, Endotrypanum/Leishmania*, and *Phytomonas* lineages (marked by arrows in Fig. 4). This proposal corresponds most closely to one advanced by Molyneux (9), who emphasized characters related to the insect vectors. As both scenarios require four independent events involving digenetic parasitism (three gains plus one loss vs. four gains), it is not possible to discriminate with these data alone. Significantly, under *any* scheme multiple origins of digenetic parasitism must be postulated. As more sequences from other taxa become available in the future, especially ones branching from the trypanosomal lineages, the proposal may be refined. As discussed below, other considerations lead us to favor the second scenario.

Relative Timing of Divergence Events within the Trypanosomatidae. One consequence of the view that digenetic parasitism has occurred multiple times is that one can now no longer associate the divergence of the parasites with that of the host with any certainty, which precludes the use of host-based divergence times, such as those applied previously by Lake *et al.* (11). To provide some perspective on the magnitude of divergence times among the Kinetoplastid genera, SSU rRNA sequence divergences within this group were compared with those among organisms for which there are some fossil estimates of divergence times (Fig. 4). Such comparisons provide at least an approximate measure of divergence times, as relative rate tests between prokaryotes and various eukaryotes, including kinetoplastids, suggest that most eukaryotic nuclear SSU rRNA genes are not evolving at widely varying rates (22). Additionally, comparisons of the divergence of total nuclear DNA in *Endotrypanum* (whose evolution is tightly linked to South American fauna) suggest that the evolutionary rate is comparable to that in metazoa (42).

SSU rRNA divergences within the Kinetoplastids are large relative to those seen in metazoans. The divergence between

Bodo and other trypanosomatids approach that observed between vertebrates and invertebrates or between plants, animals, and fungi (13.1% vs. 13.8 or 14.7%; Fig. 4). Divergence times among metazoan kingdoms and vertebrates have been estimated as 900 and 680 million years ago, respectively (refs. 44 and 45 and the references therein). SSU rRNA distances between trypanosomes and the *Leishmania/Crithidia* lineage exceed that among mammals and amphibians (6.5–9.7% vs. 3.2%; Fig. 4); the mammal amphibian divergence has been estimated as 340 million years (43, 44). Old and New World species complexes of *Leishmania* differ by 0.3%, which may reflect the separation of South America and Africa 90 million years ago (45). This biogeographical time point is supported by the observation that for the protein dihydrofolate reductase, the divergence between New and Old World *Leishmania* (8.9%; ref. 46) is comparable to that among mammalian orders (11.8%), which separated 85 million years ago.

Although it is impossible to accurately date the divergence times among trypanosomatids from these limited data, the clear implication is that many trypanosomatid divergences greatly predate the origins of both the current insect vectors (30–60 million years for the tsetse fly vector of *T. brucei*; ref. 47) and placental mammalian hosts (<85 million years). This conclusion strongly favors the second scenario above, as it rules out the ancient acquisition of digenetic parasitism required in the first scheme because the hosts were not extant.

The proposed antiquity of trypanosomatids raises many questions about the life-styles and hosts of the ancestral species. Modern trypanosomatids parasitize many invertebrate phyla and even other protozoans (1, 2), and it seems possible that many ancestral metazoans could likewise serve as suitable hosts. Because monogenetic–digenetic transitions may have occurred more often than previously suspected, it is much more difficult to unambiguously ascertain the actual ancestral host. As we incorporate studies of additional trypanosomatid taxa into these evolutionary studies, resolution of many of these questions may be forthcoming.

Our revised view of trypanosomatid evolution suggests additional perspectives on other aspects of parasite biology. For example, RNA editing of mitochondrial genes in the earliest trypanosomatid lineage (*T. brucei*) is much more extensive than in the *Crithida/Leishmania* lineage, and the patterns of editing are more similar in *Crithidia* and *Leishmania* (3, 48). This result suggests that extensive RNA editing is the ancient primitive state (see also ref. 49). Another implication concerns the origins of virulence genes required to mediate the digenetic life-style of vertebrate parasitism. Given that vertebrate parasitism has originated multiple times, it seems likely that the number of genetic alterations required to initiate such a shift could be relatively small, especially for the relatively recent *Leishmania/Endotrypanum* lineage. The events required to initiate vertebrate parasitism may thus be more amenable to study by genetic approaches than has been previously supposed.

We thank M. Sogin for assistance with computer alignments, L. Simpson for providing mitochondrial rRNA sequence alignments, our colleagues who provided strains, D. E. Dobson, G. Grimaldi, M. Petrillo-Peixoto, and I. Roitman; E. Zimmer and J. DeSalvo helped in preliminary RNA-based sequencing studies. This work was supported by grants from the National Science Foundation (S.M.B.) and Comissão de Aperfeiçoamento de Pessoal de Nível Superior and Conselho Nacional de desenvolvimento científico e technológico Brazil (CINPq) (A.P.F.), and the Harvard Medical School MacArthur Program in Molecular Parasitology and Markey Program in Developmental Biology. S.M.B. was a Burroughs Wellcome Scholar in Molecular Parasitology.

1. Vickerman, K. (1976) in *Biology of the Kinetoplastida*, eds. Lumsden, W. H. R. & Evans, D. A. (Academic, London), Vol. 1, pp. 1–76.
2. Molyneux, D. H. & Ashford, R. W. (1983) *The Biology of* Trypanosoma *and* Leishmania: *Parasites of Man and Domestic Animals* (Taylor and Francis, London).
3. Simpson, L. & Shaw, J. (1989) *Cell* **57**, 355–366.
4. Cross, G. A. M. (1990) *Annu. Rev. Immunol.* **8**, 83–110.
5. Borst, P. (1986) *Annu. Rev. Biochem.* **55**, 701–732.
6. Lainson, R. & Shaw, J. J. (1987) in *The Leishmaniasis*, eds. Killick-Kendrick, R. & Peters, W. (Academic, London), pp. 1–120.
7. Baker, J. R. (1965) in *Evolution of Parasites*, ed. Taylor, A. E. R. (Br. Soc. Parasitol., London), pp. 1–27.
8. Hoare, C. (1972) *The Trypanosomes of Mammals: A Zoological Monograph* (Blackwell, London), pp. 81–120.
9. Molyneux, D. H. (1986) in *Leishmania: Taxonomy and Phylogeny*, ed. Rioux, J. A. (IMEEE, Montpelier, France), pp. 231–240.
10. Wallace, F. G. (1966) *Exp. Parasitol.* **18**, 124–193.
11. Lake, J. A., Cruz, V. F. d. l, Ferreira, P. C. G., Morel, C. & Simpson, L. (1988) *Proc. Natl. Acad. Sci. USA* **85**, 4779–4783.
12. Hernandez, R., Rios, P., Valdes, A. M. & Pinero, D. (1990) *Mol. Biochem. Parasitol.* **41**, 207–212.
13. Gomez, E., Valdes, A. M., Pinero, D. & Hernandez, R. (1991) *Mol. Biol. Evol.* **8**, 254–259.
14. Briones, M. R. S., Nelson, K., Beverley, S. M., Affonso, H. T., Camargo, E. P. & Floeter-Winter, L. M. (1992) *Mol. Biochem. Parasitol.* **53**, 121–128.
15. Wiley, R. L., Walne, P. L. & Kivic, P. (1988) *Crit. Rev. Plant Sci.* **7**, 303–340.
16. Hajduk, S. L., Siqueira, A. M. & Vickerman, K. (1986) *Mol. Cell. Biol.* **6**, 4372–4378.
17. Attias, M. & de Souza, W. (1986) *J. Protozool.* **33**, 84–87.
18. Beverley, S. M. (1988) *Nucleic Acids Res.* **16**, 925–938.
19. Bingham, P. M., Levis, R. & Reubin, G. M. (1981) *Cell* **25**, 693–704.
20. Brown, P. C., Beverley, S. M. & Schimke, R. T. (1981) *Mol. Cell. Biol.* **1**, 1077–1083.
21. Looker, D., Miller, L. A., Elwood, H. J., Stickel, S. & Sogin, M. L. (1988) *Nucleic Acids Res.* **16**, 7198–7198.
22. Sogin, M. L., Elwood, H. J. & Gunderson, J. H. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 1383–1387.
23. Schnare, M. N., Collings, J. C. & Gray, M. W. (1986) *Curr. Genet.* **10**, 405–410.
24. Spencer, D. F., Collings, J. C., Schnare, M. N. & Gray, M. W. (1987) *EMBO J.* **6**, 1063–1071.
25. Campbell, D. A., Kubo, K., Clark, C. G. & Boothroyd, J. C. (1987) *J. Mol. Biol.* **196**, 113–124.
26. Saiki, R. K., Gelfand, D. H., Stoffel, S., Scharf, S. J., Higuchi, R., Horn, G. T., Mullis, K. B. & Erlich, H. A. (1988) *Science* **239**, 487–494.
27. Messing, J. (1983) *Methods Enzymol.* **101**, 20–78.
28. Neefs, J. M., Van de Peer, Y., Hendriks, L. & De Wachter, R. (1990) *Nucleic Acids Res.* **18**, S2237–S2317.
29. Elwood, H. J., Holsen, G. J. & Sogin, M. L. (1985) *Mol. Biol. Evol.* **2**, 399–410.
30. Swofford, D. L. (1991) *PAUP* (Illinois Natural History Survey, Champagne, IL), Version 3.1.1.
31. Uliana, S. R. B., Nelson, K., Beverley, S. M., Camargo, E. P. & Floeter-Winter, L. M. (1993) *J. Protozool.*, in press.
32. Arruda, D. V. M., Reinach, F. C., Colli, W. & Zingales, B. (1990) *Mol. Biochem. Parasitol.* **40**, 35–42.
33. Tibayrenc, M., Ward, P., Moya, A. & Ayala, F. J. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 115–119.
34. van Eys, G. J. J. M., Schoone, G. J., Kroon, N. C. M. & Ebeling, S. B. (1992) *Mol. Biochem. Parasitol.* **51**, 133–142.
35. Beverley, S. M., Ismach, R. B. & McMahon-Pratt, D. (1987) *Proc. Natl. Acad. Sci. USA* **84**, 484–488.
36. Beverley, S. M., Ellenberger, T. E. & Cordingley, J. S. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 2584–2588.
37. Hughes, D., Shonekan, O. A. & Simpson, L. (1989) *Mol. Biochem. Parasitol.* **34**, 155–166.
38. Simpson, L., Neckelmann, N., de la Cruz, V., Simpson, A., Feagin, J., Jasmer, D. & Stuart, K. (1987) *J. Biol. Chem.* **262**, 6182–6196.
39. Alonso, G., Guevara, P. & Ramirez, J. L. (1992) *Mem. Inst. Oswaldo Cruz* **87**, 517–523.
40. Langford, C. K., Ullman, B. & Landfear, S. M. (1992) *Exp. Parasitol.* **74**, 360–361.
41. Parsons, M., Stuart, K. & Smiley, B. L. (1991) *Mol. Biochem. Parasitol.* **73**, 101–105.
42. Lopes, A. H. d. C. S., Iovannisci, D., Petrillo-Peixoto, M., McMahon-Pratt, D. & Beverley, S. M. (1990) *Mol. Biochem. Parasitol.* **40**, 151–162.
43. Baba, M. L., Darga, L. L., Goodman, M. & Czelusniak, J. (1981) *J. Mol. Evol.* **17**, 197–213.
44. Britten, R. J. (1986) *Science* **231**, 1393–1398.
45. Tarling, D. H. (1980) in *Evolutionary Biology of the New World Monkeys and Continental Drift*, eds. Ciochon, R. L. & Charelli, A. B. (Plenum, New York), pp. 1–41.
46. Nelson, K., Alonso, G., Langer, P. J. & Beverley, S. M. (1990) *Nucleic Acids Res.* **18**, 2819–2819.
47. Lambrecht, F. (1980) *Proc. Am. Philos. Soc.* **124**, 367–384.
48. Benne, R. (1989) *Biochim. Biophys. Acta* **1007**, 131–139.
49. Landweber, L. F. & Gilbert, W. (1994) *Proc. Natl. Acad. Sci. USA*, in press.